SALES ANALYSIS

In [5]:
```
import pandas as pd
import io

df = pd.read_csv('Sales Data.csv')
print(df)
```

```
        Unnamed: 0  Order ID                 Product  Quantity Ordered  \
0                0    295665      Macbook Pro Laptop                 1   
1                1    295666      LG Washing Machine                 1   
2                2    295667    USB-C Charging Cable                 1   
3                3    295668        27in FHD Monitor                 1   
4                4    295669    USB-C Charging Cable                 1   
...            ...       ...                     ...               ...   
185945       13617    222905  AAA Batteries (4-pack)                 1   
185946       13618    222906        27in FHD Monitor                 1   
185947       13619    222907    USB-C Charging Cable                 1   
185948       13620    222908    USB-C Charging Cable                 1   
185949       13621    222909  AAA Batteries (4-pack)                 1   

        Price Each           Order Date  \
0          1700.00  2019-12-30 00:01:00   
1           600.00  2019-12-29 07:03:00   
2            11.95  2019-12-12 18:21:00   
3           149.99  2019-12-22 15:13:00   
4            11.95  2019-12-18 12:38:00   
...            ...                  ...   
185945        2.99  2019-06-07 19:02:00   
185946      149.99  2019-06-01 19:29:00   
185947       11.95  2019-06-22 18:57:00   
185948       11.95  2019-06-26 18:35:00   
185949        2.99  2019-06-25 14:33:00   

                              Purchase Address  Month    Sales  \
0       136 Church St, New York City, NY 10001     12  1700.00   
1          562 2nd St, New York City, NY 10001     12   600.00   
2         277 Main St, New York City, NY 10001     12    11.95   
3          410 6th St, San Francisco, CA 94016     12   149.99   
4                43 Hill St, Atlanta, GA 30301     12    11.95   
...                                        ...    ...      ...   
185945           795 Pine St, Boston, MA 02215      6     2.99   
185946   495 North St, New York City, NY 10001      6   149.99   
185947   319 Ridge St, San Francisco, CA 94016      6    11.95   
185948    916 Main St, San Francisco, CA 94016      6    11.95   
185949          209 11th St, Atlanta, GA 30301      6     2.99   

                 City  Hour  
0       New York City     0  
1       New York City     7  
2       New York City    18  
3       San Francisco    15  
4             Atlanta    12  
...               ...   ...  
185945         Boston    19  
185946  New York City    19  
185947  San Francisco    18  
185948  San Francisco    18  
185949        Atlanta    14  

[185950 rows x 11 columns]
```

In [7]:
```
#import all required packages..
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [8]:

```
#loading sales analysis dataset..

df = pd.read_csv('Sales Data.csv',header=None)
df
```

```
C:\Users\21b01\AppData\Local\Temp\ipykernel_16100\2687814957.py:3: DtypeWarning: Columns (1,3,4,
7,8,10) have mixed types. Specify dtype option on import or set low_memory=False.
  df = pd.read_csv('Sales Data.csv',header=None)
```

Out[8]:

	0	1	2	3	4	5	6	7	8	9	10
0	NaN	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	Hour
1	0.0	295665	Macbook Pro Laptop	1	1700.0	2019-12-30 00:01:00	136 Church St, New York City, NY 10001	12	1700.0	New York City	0
2	1.0	295666	LG Washing Machine	1	600.0	2019-12-29 07:03:00	562 2nd St, New York City, NY 10001	12	600.0	New York City	7
3	2.0	295667	USB-C Charging Cable	1	11.95	2019-12-12 18:21:00	277 Main St, New York City, NY 10001	12	11.95	New York City	18
4	3.0	295668	27in FHD Monitor	1	149.99	2019-12-22 15:13:00	410 6th St, San Francisco, CA 94016	12	149.99	San Francisco	15
...	...	...	...	...	...	...	...	...	...	...	...
185946	13617.0	222905	AAA Batteries (4-pack)	1	2.99	2019-06-07 19:02:00	795 Pine St, Boston, MA 02215	6	2.99	Boston	19
185947	13618.0	222906	27in FHD Monitor	1	149.99	2019-06-01 19:29:00	495 North St, New York City, NY 10001	6	149.99	New York City	19
185948	13619.0	222907	USB-C Charging Cable	1	11.95	2019-06-22 18:57:00	319 Ridge St, San Francisco, CA 94016	6	11.95	San Francisco	18
185949	13620.0	222908	USB-C Charging Cable	1	11.95	2019-06-26 18:35:00	916 Main St, San Francisco, CA 94016	6	11.95	San Francisco	18
185950	13621.0	222909	AAA Batteries (4-pack)	1	2.99	2019-06-25 14:33:00	209 11th St, Atlanta, GA 30301	6	2.99	Atlanta	14

185951 rows × 11 columns

In [9]:

```
df.head()
```

Out[9]:

	0	1	2	3	4	5	6	7	8	9	10
0	NaN	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	Hour
1	0.0	295665	Macbook Pro Laptop	1	1700.0	2019-12-30 00:01:00	136 Church St, New York City, NY 10001	12	1700.0	New York City	0
2	1.0	295666	LG Washing Machine	1	600.0	2019-12-29 07:03:00	562 2nd St, New York City, NY 10001	12	600.0	New York City	7
3	2.0	295667	USB-C Charging Cable	1	11.95	2019-12-12 18:21:00	277 Main St, New York City, NY 10001	12	11.95	New York City	18
4	3.0	295668	27in FHD Monitor	1	149.99	2019-12-22 15:13:00	410 6th St, San Francisco, CA 94016	12	149.99	San Francisco	15

In [10]:

```
#replacing empty value with 0.
df.fillna(0,inplace=True)
```

In [11]:
```
df.head()
```

Out[11]:

	0	1	2	3	4	5	6	7	8	9	10
0	0.0	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	Hour
1	0.0	295665	Macbook Pro Laptop	1	1700.0	2019-12-30 00:01:00	136 Church St, New York City, NY 10001	12	1700.0	New York City	0
2	1.0	295666	LG Washing Machine	1	600.0	2019-12-29 07:03:00	562 2nd St, New York City, NY 10001	12	600.0	New York City	7
3	2.0	295667	USB-C Charging Cable	1	11.95	2019-12-12 18:21:00	277 Main St, New York City, NY 10001	12	11.95	New York City	18
4	3.0	295668	27in FHD Monitor	1	149.99	2019-12-22 15:13:00	410 6th St, San Francisco, CA 94016	12	149.99	San Francisco	15

In [16]:
```
data = pd.read_csv("Sales Data.csv").dropna()
data.head()
```

Out[16]:

	Unnamed: 0	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	Hour
0	0	295665	Macbook Pro Laptop	1	1700.00	2019-12-30 00:01:00	136 Church St, New York City, NY 10001	12	1700.00	New York City	0
1	1	295666	LG Washing Machine	1	600.00	2019-12-29 07:03:00	562 2nd St, New York City, NY 10001	12	600.00	New York City	7
2	2	295667	USB-C Charging Cable	1	11.95	2019-12-12 18:21:00	277 Main St, New York City, NY 10001	12	11.95	New York City	18
3	3	295668	27in FHD Monitor	1	149.99	2019-12-22 15:13:00	410 6th St, San Francisco, CA 94016	12	149.99	San Francisco	15
4	4	295669	USB-C Charging Cable	1	11.95	2019-12-18 12:38:00	43 Hill St, Atlanta, GA 30301	12	11.95	Atlanta	12

In [17]:
```
products = data["Product"].unique()
```

In [ ]:
```
#SOLVING TASKS
```

In [18]:
```
product_quantities = data[["Product", "Quantity Ordered"]].groupby("Product").sum()
product_quantities.head()
```

Out[18]:

	Quantity Ordered
Product	
20in Monitor	4129
27in 4K Gaming Monitor	6244
27in FHD Monitor	7550
34in Ultrawide Monitor	6199
AA Batteries (4-pack)	27635

In [19]:
```
product_quantities = product_quantities["Quantity Ordered"].sort_values(ascending=False)
product_quantities.head()
```

Out[19]:
```
Product
AAA Batteries (4-pack)      31017
AA Batteries (4-pack)       27635
USB-C Charging Cable        23975
Lightning Charging Cable    23217
Wired Headphones            20557
Name: Quantity Ordered, dtype: int64
```

```
In [ ]: #5 most sold products in terms of number of quantity orders
```

```
In [20]: my_criterion = { product : product_quantities[product]/data[data["Product"]==product]["Price Each"
         my_criterion = pd.Series(my_criterion).sort_values(ascending=False)
         my_criterion.head()
```

```
Out[20]: AAA Batteries (4-pack)      10373.578595
         AA Batteries (4-pack)        7196.614583
         USB-C Charging Cable         2006.276151
         Wired Headphones             1714.512093
         Lightning Charging Cable     1552.976589
         dtype: float64
```

```
In [ ]: #Searching what products sold the most in every city.
```

```
In [21]: cities = data["City"].unique()
         print(cities)
```

```
         [' New York City' ' San Francisco' ' Atlanta' ' Portland' ' Dallas'
          ' Los Angeles' ' Boston' ' Austin' ' Seattle']
```

```
In [22]: product_sells_per_sity = {
             city : data[data["City"] == city][["Product", "Quantity Ordered"]].groupby("Product").sum()["Q
         }
```

In [23]:

```
for city, products in product_sells_per_sity.items():
    print(city)
    print(products.head(), end="\n\n")
```

```
 New York City
Product
AAA Batteries (4-pack)      4124
AA Batteries (4-pack)       3630
USB-C Charging Cable        3269
Lightning Charging Cable    3041
Wired Headphones            2707
Name: Quantity Ordered, dtype: int64

 San Francisco
Product
AAA Batteries (4-pack)      7408
AA Batteries (4-pack)       6555
USB-C Charging Cable        5894
Lightning Charging Cable    5557
Wired Headphones            4966
Name: Quantity Ordered, dtype: int64

 Atlanta
Product
AAA Batteries (4-pack)      2359
AA Batteries (4-pack)       2193
USB-C Charging Cable        1915
Lightning Charging Cable    1879
Wired Headphones            1579
Name: Quantity Ordered, dtype: int64

 Portland
Product
AAA Batteries (4-pack)      2080
AA Batteries (4-pack)       1939
USB-C Charging Cable        1582
Lightning Charging Cable    1531
Wired Headphones            1362
Name: Quantity Ordered, dtype: int64

 Dallas
Product
AAA Batteries (4-pack)      2504
AA Batteries (4-pack)       2261
Lightning Charging Cable    1864
USB-C Charging Cable        1852
Wired Headphones            1669
Name: Quantity Ordered, dtype: int64

 Los Angeles
Product
AAA Batteries (4-pack)      4967
AA Batteries (4-pack)       4438
USB-C Charging Cable        3782
Lightning Charging Cable    3772
Wired Headphones            3270
Name: Quantity Ordered, dtype: int64

 Boston
Product
AAA Batteries (4-pack)      3461
AA Batteries (4-pack)       3016
USB-C Charging Cable        2561
Lightning Charging Cable    2491
Wired Headphones            2222
Name: Quantity Ordered, dtype: int64

 Austin
Product
AAA Batteries (4-pack)      1668
AA Batteries (4-pack)       1424
Lightning Charging Cable    1307
USB-C Charging Cable        1251
Wired Headphones            1130
Name: Quantity Ordered, dtype: int64

 Seattle
Product
AAA Batteries (4-pack)      2446
AA Batteries (4-pack)       2179
```

```
USB-C Charging Cable        1869
Lightning Charging Cable    1775
Wired Headphones            1652
Name: Quantity Ordered, dtype: int64
```

In [ ]:

```
#Similar tendency in most sold products in every city.
```

In [24]:

```
my_criterion = { product : data[data["Product"]==product]["Quantity Ordered"].mean() for product i
my_criterion = pd.Series(my_criterion).sort_values(ascending=False)
my_criterion.head()
```

Out[24]:

```
2446   NaN
2179   NaN
1869   NaN
1775   NaN
1652   NaN
dtype: float64
```

In [25]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 185950 entries, 0 to 185949
Data columns (total 11 columns):
 #   Column            Non-Null Count   Dtype  
---  ------            --------------   -----  
 0   Unnamed: 0        185950 non-null  int64  
 1   Order ID          185950 non-null  int64  
 2   Product           185950 non-null  object 
 3   Quantity Ordered  185950 non-null  int64  
 4   Price Each        185950 non-null  float64
 5   Order Date        185950 non-null  object 
 6   Purchase Address  185950 non-null  object 
 7   Month             185950 non-null  int64  
 8   Sales             185950 non-null  float64
 9   City              185950 non-null  object 
 10  Hour              185950 non-null  int64  
dtypes: float64(2), int64(5), object(4)
memory usage: 15.6+ MB
```

In [26]:

```
data.describe()
```

Out[26]:

	Unnamed: 0	Order ID	Quantity Ordered	Price Each	Month	Sales	Hour
count	185950.000000	185950.000000	185950.000000	185950.000000	185950.000000	185950.000000	185950.000000
mean	8340.388475	230417.569379	1.124383	184.399735	7.059140	185.490917	14.413305
std	5450.554093	51512.737110	0.442793	332.731330	3.502996	332.919771	5.423416
min	0.000000	141234.000000	1.000000	2.990000	1.000000	2.990000	0.000000
25%	3894.000000	185831.250000	1.000000	11.950000	4.000000	11.950000	11.000000
50%	7786.000000	230367.500000	1.000000	14.950000	7.000000	14.950000	15.000000
75%	11872.000000	275035.750000	1.000000	150.000000	10.000000	150.000000	19.000000
max	25116.000000	319670.000000	9.000000	1700.000000	12.000000	3400.000000	23.000000

In [ ]: